

Reflections on the First Man vs. Machine No-Limit Texas Hold 'em Competition

Sam Ganzfried

Assistant Professor, Computer Science, Florida International University, Miami FL

PhD, Computer Science Department, Carnegie Mellon University, 2015

sam.ganzfried@gmail.com

Brains vs. Artificial Intelligence

- April 24-May 8, 2015 at Rivers Casino in Pittsburgh, PA
 - The competition was organized by Carnegie Mellon University Professor Tuomas Sandholm. Collaborators were Tuomas Sandholm and Noam Brown.
- 20,000 hands of two-player no-limit Texas hold ‘em between “Claudico” and Dong Kim, Jason Les, Bjorn Li, Doug Polk
 - 80,000 hands in total
 - Two 750-hand sessions per day

“Duplicate” scoring

- Suppose Dong has pocket aces and Claudico has pocket kings and Dong wins \$5,000. Did Dong “outplay” Claudico?
 - What if Bjorn had pocket kings against Claudico’s pocket aces in the same situation (same public cards dealt on the board), and Claudico won \$10,000?



Brains



Doug Polk
@DougPolkPoker

March HUNL PR
1 West Coast Gangsters
2 Big Dick
3 AZNflushie (RIP)
4 Rumble man
5 Swarmmy
6 Kaby
7 Ike
8 wheyprotein
9 80%carry
10 muumi

RETWEETS 4
LIKES 13



4:24 AM - 28 Feb 2015



Nick Frame
@TCfromUB

Follow

The REAL power rankings for OCT 2014 are out

TC power rankings OCT 2014

1. WCG (0)
2. ike (+1)
3. sauce (+1)
4. TCfromUB (+1)
5. jungle (+5)
6. pandorasbux (-4)
7. kabydf (0)
8. donger (-2)
9. carrycakes (-1)
10. KPR (-1)
11. asianflushie (+3)
12. kanu7 (+3)
13. bajskorven (U)
14. OTBredbaron (U)
15. Rperfumo (-4)
16. mokoma1 (0)
17. Billiomucks (-5)
18. dougiedan (-5)
19. ForTheSwarm (U)
20. Willhasha (U)

Brains

02-04-2015, 09:53 AM

#1

Donger Kim

enthusiast



Join Date: Feb 2015
Location: South Korea
@dongerkim
Posts: 54

Donger Kim to Nick Frame (TCfromUB) HU Challenge

I am a high-stakes heads up nlhe regular on PokerStars where I play under the name "Donger Kim". There's been quite a bit of discussion on heads-up rankings lately, particularly from TCfromUB (Nick Frame, TooCuriosso1 on 2p2). I've played quite a bit with him and think he's a top player. I respect his game and it would be humbling to play him and represent my country.

However, as he ranks himself ahead of me, I'd like to have a chance to play him in a challenge-type format. I think it would be a fun experience and something that would also be enjoyable for the community.

I propose we do a 15k hand challenge at 100/200 nl with a \$50k sidebet escrowed with ike or sauce. I suggest we put some reasonable time frame conditions on this, we're both grinders so we should be able to finish this in a 1-2 week time frame.

Nick, let me know when you'd like to begin. Ideally, I'd like to get started right away.

Brains

MAR
2

Donger Kim wins heads-up challenge against TCfromUB

[Poker News](#)

[G+](#) [D](#)

[f Like](#)

1

[✉](#)

[t](#)

[f](#)

[+](#)

[Share](#)

Dong "Donger Kim" Kim won \$103,992 from Nick "TCfromUB" Frame in the 15,000 hand heads-up challenge, which not only earned him the respect of the high stakes community, but also an additional \$15,000 from the sidebets for the challenge.



Results

- Humans won by 732,713 chips, which corresponds to 9.16 big blinds per 100 hands (BB/100) (SB = 50, BB = 100). Players started each hand with 200 big blinds (20,000 chips).
 - Statistically significant at 90% confidence level, but not 95%
- Dong Kim beat Nick Frame by 13.87 BB/100
 - \$103,992 over 15,000 hands with 25-50 blinds
- Doug Polk beat Ben Sulsky by 24.67 BB/100
 - \$740,000 over 15,000 hands with 100-200 blinds

Payoffs

- Prize pool of \$100,000 distributed to the humans depending on their individual profits.

If $x_1 > x_4$

$$p_1 = \$10,000 + \$60,000 \cdot \frac{x_1 - x_4}{x_1 + x_2 + x_3 - 3x_4}$$

$$p_2 = \$10,000 + \$60,000 \cdot \frac{x_2 - x_4}{x_1 + x_2 + x_3 - 3x_4}$$

$$p_3 = \$10,000 + \$60,000 \cdot \frac{x_3 - x_4}{x_1 + x_2 + x_3 - 3x_4}$$

$$p_4 = \$10,000$$

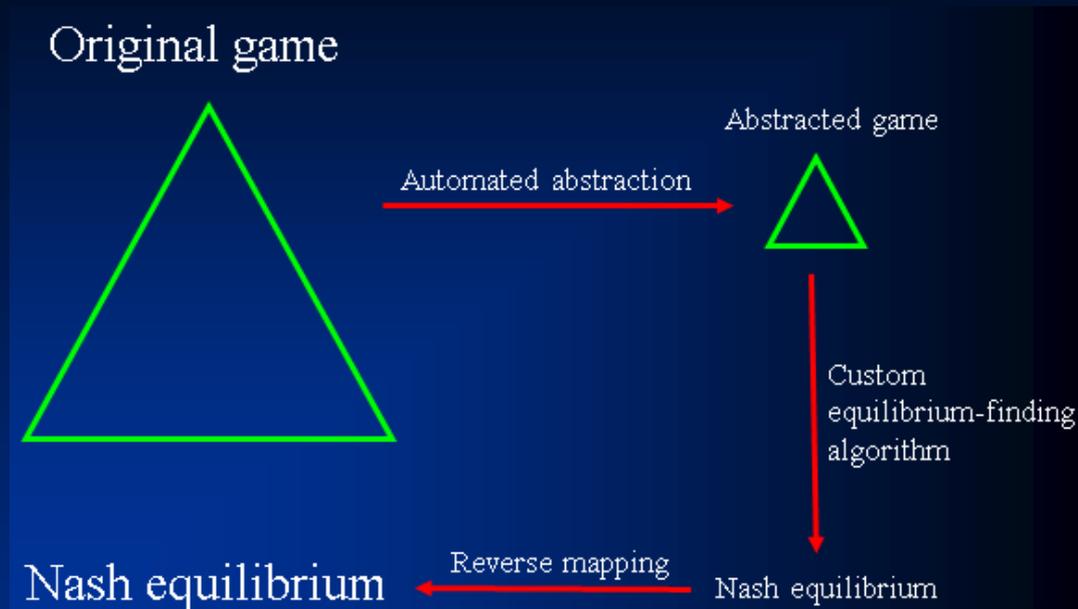
Else

$$p_1 = p_2 = p_3 = p_4 = \$25,000$$

I Limp!

- “**Limping is for Losers.** This is *the most important fundamental* in poker -- for every game, for every tournament, every stake: If you are the first player to voluntarily commit chips to the pot, open for a raise. Limping is inevitably a losing play. If you see a person at the table limping, you can be fairly sure he is a bad player. Bottom line: If your hand is worth playing, it is worth raising” [Phil Gordon’s Little Gold Book, 2011]
- Claudico limps close to 10% of its hands
 - Based on humans’ analysis it profited overall from the limps
- Claudico makes many other unconventional plays (e.g., small bets of 10% pot and all-in bets for 40 times pot)

Architecture



- Offline abstraction and equilibrium computation
 - EC used Pittsburgh's Blacklight supercomputer with 961 cores
- Action translation
- Post-processing
- Endgame solving

Abstraction and equilibrium

- Create hierarchical information abstraction that allows us to assign disjoint components of the game tree to different blades so the trajectory of each sample only accesses information sets located on the same blade.
 - First cluster public information at some early point in the game (public flop cards in poker), then cluster private information separately for each public cluster.
- Run modified version of external-sampling MCCFR
 - Samples one pair of preflop hands per iteration. For the later betting rounds, each blade samples public cards from its public cluster and performs MCCFR within each cluster.

Action translation



- $f_{A,B}(x) \equiv$ probability we map x to A
 - Will also denote as just $f(x)$

Natural approach

- If $x < \frac{A+B}{2}$, then map x to A ; otherwise, map x to B
- Called the **deterministic arithmetic** mapping

- Suppose pot is 1, stacks are 100
- Suppose we are using the {fold, call, pot, all-in} action abstraction
 - “previous expert knowledge [has] dictated that if only a single bet size [in addition to all-in] is used everywhere, it should be pot sized” [Hawkin et al., AAAI 2012]
- Suppose opponent bets x in $(1,100)$
 - So $A = 1, B = 100$

- Suppose we call a bet of 1 with probability $\frac{1}{2}$ with a medium-strength hand
- Suppose the opponent has a very strong hand
- His expected payoff of betting 1 will be:
$$(1 \cdot \frac{1}{2}) + (2 \cdot \frac{1}{2}) = 1.5$$
- If instead he bets 50, his expected payoff will be:
$$(1 \cdot \frac{1}{2}) + (51 \cdot \frac{1}{2}) = 26$$
- He gains \$24.50 by exploiting our translation mapping!
- Tartanian1 lost to an agent that didn't look at its private cards in 2007 ACPC using this mapping!

- Randomized arithmetic: map x to A with probability

$$f(x) = \frac{B-x}{B-A}$$

- Deterministic geometric: If $\frac{A}{x} > \frac{x}{B}$, map x to A ;
otherwise, map x to B

- Used by Tartanian2 in 2008

- Randomized geometric 1

- $f(x) = \frac{A(B-x)}{A(B-x) + x(x-A)}$

- Used by Alberta 2009-present

- Randomized geometric 2

- $f(x) = \frac{A(B+x)(B-x)}{(B-A)(x^2 + AB)}$

- Used by CMU 2010-2011

Pseudo-harmonic mapping

- Maps opponent's bet x to one of the nearest sizes in the abstraction A, B according to:
- $$f(x) = \frac{(B-x)(1+A)}{(B-A)(1+x)}$$
- $f(x)$ is probability that x is mapped to A
- Example: suppose opponent bets 100 into pot of 500, and closest sizes are “check” (i.e., bet 0) or to bet 0.25 pot. So $A = 0$, $x = 0.2$, $B = 0.25$.
- Plugging these in gives $f(x) = 1/6 = 0.167$.

Post-processing

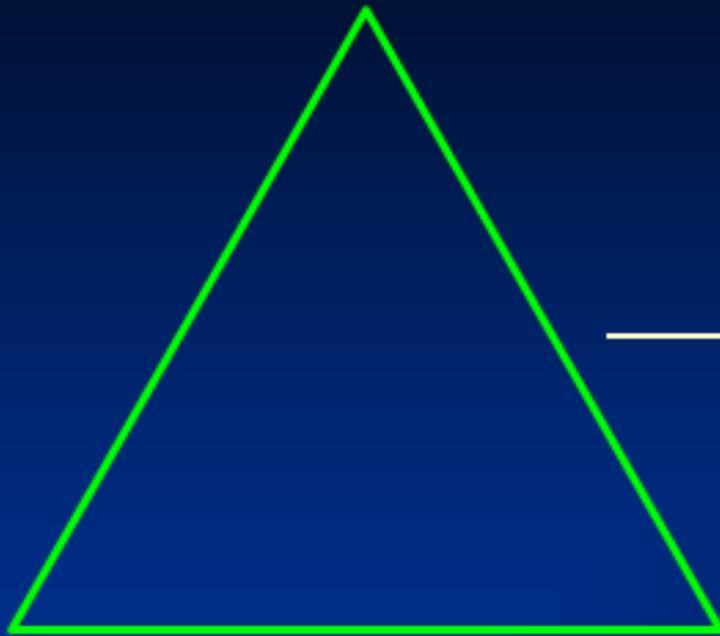
- *Thresholding*: round action probabilities below c down to 0 (then renormalize)
- *Purification* is extreme case where we play maximal-probability action with probability 1
- Generalizations:
 - Bundle similar actions
 - Add preference for conservative actions
- First separate actions into {fold, call, “bet”}
 - If probability of folding exceeds a threshold parameter, fold with prob. 1
 - Else, follow purification between fold, call, and “meta-action” of “bet.”
 - If “bet” is selected, then follow purification within the specific bet actions.

| Threshold | Exploitability of GS6 | Exploitability of Hyperborean |
|-----------|-----------------------|-------------------------------|
| None | 463.591 | 235.209 |
| 0.05 | 326.119 | 243.705 |
| 0.15 | 318.465 | 258.53 |
| 0.25 | 335.048 | 277.841 |
| Purified | 349.873 | 437.242 |

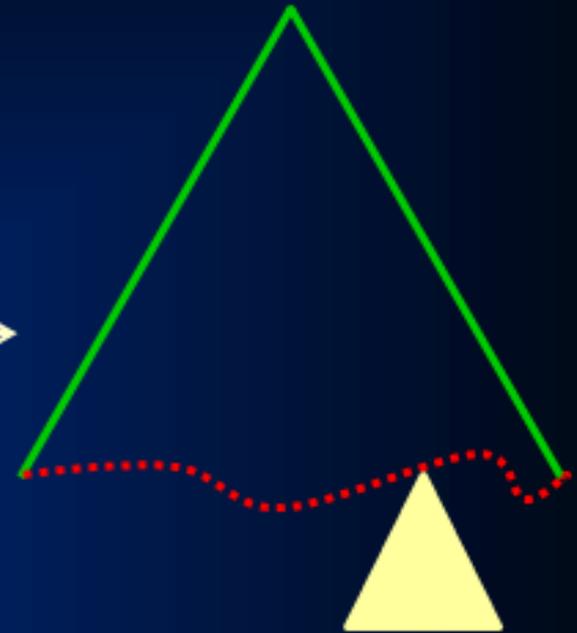
Table 4: Results for full-game worst-case exploitabilities of several strategies in two-player limit Texas Hold'em. Results are in milli big blinds per hand. Bolded values indicate the lowest exploitability achieved for each strategy.

| | Hyperborean.iro | Slumbot | Average | Min |
|-------------------|------------------------|----------------|----------------|------------|
| No Thresholding | +30 ± 32 | +10 ± 27 | +20 | +10 |
| Purification | +55 ± 27 | +19 ± 22 | +37 | +19 |
| Thresholding-0.15 | +35 ± 30 | +19 ± 25 | +27 | +19 |
| New-0.2 | +39 ± 26 | +103 ± 21 | +71 | +39 |

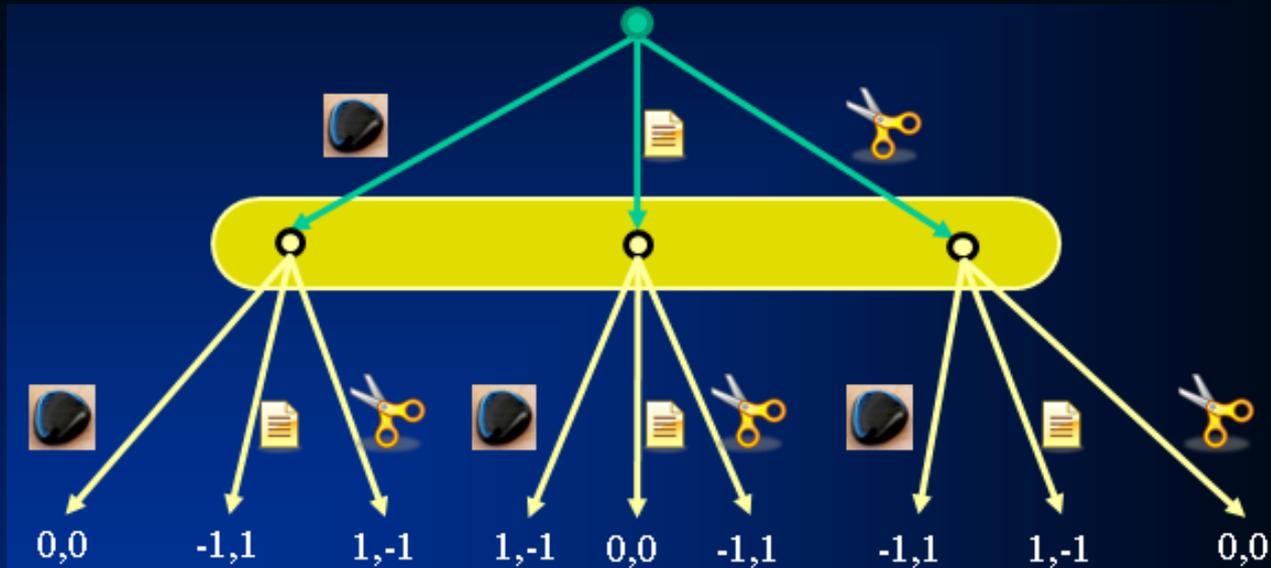
Endgame solving



Strategies for entire game
computed offline



Endgame strategies
computed in real time to
greater degree of accuracy



- Developed new efficient algorithm for endgame solving that requires only $O(n)$ instead of $O(n^2)$ strategy table lookups
- Our approach improved performance against strongest 2013 ACPC agents: 87+-50 vs. Hyperborean and 29+-25 vs. Slumbot
- Doug Polk related to me in personal communication after the competition that he thought the river strategy of Claudico using the endgame solver was the strongest part of the agent.

Problematic hands

1. We had A4s and folded preflop after putting in over half of our stack (human had 99).
 - We only need to win 25% of time against opponent's distribution for call to be profitable (we win 33% of time against 99).
 - Translation mapped opponent's raise to smaller size, which caused us to look up strategy computed thinking that pot size was much smaller than it was (7,000 vs. 10,000)
2. We had KT and folded to an all-in bet on turn after putting in $\frac{3}{4}$ of our stack despite having top pair and a flush draw
 - Human raised slightly below smallest size in our abstraction and we interpreted it as a call
 - Both 1 and 2 due to “off-tree problem” (endgame solving solves this)
3. Large all-in bet of 19,000 into small pot of 1700 on river without “blocker”
 - E.g., 3s2c better all-in bluff hand than 3c2c on JsTs4sKcQh
 - Endgame abstraction algorithm doesn't fully account for “card removal”

Equilibrium vs. “learning”

- Garry Kasparov discusses “freestyle” chess tournament
- “The winner was revealed to be not a grandmaster with a state-of-the-art PC but a pair of amateur American chess players using three computers at the same time.”

Human “learning”

- Modify own play over course of hands within session, and between different sessions
- Analyze database of Claudico’s play at night
 - Personal data analyst
- Discuss hands in real time with other humans

“Brains”

Libra VI 20-50 bb, fast - \$0.10/\$0.25 USD - No Limit Hold'em

POT \$1.55

Ectrie Sitting Out Tsepaka Sitting Out

ziapeg \$5

Hand: J♥ J♥ 7♥ 7♥ 6♠ 6♠ 4♥ 4♥

sobakadoter \$5 jose fco fde \$4.30

sobakadoter/(5)? jose fco fde/(5)?

40%/40.0%/∞/\$0 0%/0.0%/0.0/\$-1

Buttons: Preflop, Flop, Turn, River

| Player | Total | Blinds | EP | MP | LP | Cupii |
|----------------------------------|-------|--------|----|------|----|-------|
| sobakadoter | 40% | 50% | 0% | 100% | 0% | |
| VP\$IP | 40% | 50% | 0% | 100% | 0% | |
| PreFlop Raise | 40% | 50% | 0% | 100% | 0% | |
| Aggr Factor | ∞ | ∞ | 0 | ∞ | 0 | |
| Fold Big Blind to Steal | 100% | | | | | |
| Attempt to Steal | 100% | | | | | |
| Fold to Continuation Bet on Flop | 0% | | | | | |
| Won Money at Show Down | 0% | | | | | |
| BB/100 requested TIME | 28 | | | | | |

Chat Notes Stats

Dealer: sobakadoter, it's act

Dealer: Player sobakadoter requested TIME

Dealer: Game #53879391411: hamag27 wins pot [\$0.20]

Join Waiting List

There are no players in line

“AI”



AI “learning”?

- Equilibrium computation
- Multiple strategies
- Switching action translation mapping
 - E.g., from randomized to deterministic
- Degree of thresholding in each round
- Endgame solver
 - Whether to use at all
 - Granularity of endgame (size of action and information abstraction)
 - Which bet sizes to include

- Science vs. entertainment
 - Is it ok for brains to utilize “AI” and AI to utilize “brains?” Or do we want strictly Brains vs AI?
- Can we decrease variance further?
 - Also used “all-in EV”
- Are “hybrid” human/AI agents future of AI? Or does the field want to stick to purely algorithmic approaches (at expense of performance)
- “Flexible” algorithms
 - parameters that can be tuned in real-time by human expert

Conclusions and directions

- Two most important avenues for improvement
 - Solving the “off-tree problem”
 - Improved approach for information abstraction that better accounts for card removal/“blockers”
- Improved theoretical understanding of endgame solving
 - Works very well in practice despite lack of guarantees
 - Newer decomposition approach with guarantees does worse
- Bridge abstraction gap
 - Approaches with guarantees only scale to small games
- Diverse applications of equilibrium computation
- Action translation axioms
- Theoretical understanding of post-processing success

- www.ganzfriedresearch.com
- <http://forumserver.twoplustwo.com/29/news-views-gossip-sponsored-online-poker-report/wcgrider-dong-kim-jason-les-bjorn-li-play-against-new-hu-bot-1526750/>
- <https://www.youtube.com/watch?v=phRAyF1rq0I>

The screenshot shows a poker hand in progress. The title is "BRAINS VS. ARTIFICIAL INTELLIGENCE". The player "Doug Polk" is on the left, and the AI "Claudio Polk" is on the right. The community cards are A♥, 3♥, A♦, 8♠, 9♠. Doug Polk's hand is K♥, 5♥. The pot is 2624. The AI's action is r250CB100x625CMb437Ck. The interface includes a video feed of Doug Polk, a balance table, and various logos like Carnegie Mellon University, Microsoft, and Rivers Casino.

Player Balance Table:

| Player | Balance |
|--------------|-----------|
| dougpolk2 | 364 / 750 |
| Claudio_Polk | -18179 |
| DougPolk | 18179 |

Community Cards: A♥, 3♥, A♦, 8♠, 9♠

Doug Polk Hand: K♥, 5♥

Claudio Polk Hand: (Hidden)

Pot: 2624

AI Action: r250CB100x625CMb437Ck

Wager Amount: 100

Buttons: Fold, Check, Bet In